Journal of Academia Vol. 9, Issue 2 (2021) 154 – 165

SCREENING OF MEDIUM COMPOSITIONS FOR RECOMBINANT LIPASE PRODUCTION VIA TWO-LEVEL FRACTIONAL FACTORIAL DESIGN

Aina Yasmin Daud[1], Wan Siti Atikah Wan Omar[2*]

[1]*Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor Darul Ehsan, Malaysia,*
[2]*Universiti Teknologi MARA (UiTM), Cawangan Pahang, Kampus Jengka, 26400 Bandar Tun Abdul Razak Jengka, Pahang, Malaysia*

[*]*Corresponding author: atikah_bio@uitm.edu.my*

Abstract

The demand for lipase in the food and detergent industries is gradually increasing with the human population. Therefore, the production of recombinant lipase is an alternative to accommodate the high supply-demand caused. Large-scale enzymatic production aims to produce a maximum yield with a minimum, medium cost. An experimental design using a two-level fractional factorial design (2^{k-1} FFD) was conducted at a laboratory scale. The experiment was carried out to measure and screen the effect of five medium components as the factors, namely ammonium sulphate ($(NH_4)_2SO_4$), glycerol, yeast extract, monopotassium dihydrogen phosphate (KH_2PO_4) and dipotassium phosphate (K_2HPO_4). The factors were exploited respectively at low and high-level parameter ranges in g/L, $(NH_4)_2SO_4$ 7.00-14.00; glycerol 17.50 – 35.00, yeast extract 10.00 – 20.00; KH_2PO_4 1.00-2.00; K_2HPO_4 8.25-16.50). The combination of experimental design with 2^{5-1} fractional factorial design reflected 16 shake flasks of submerged fermentation of recombinant *E. coli* (r*E. coli*) harbouring recombinant lipase. Responses of lipase and total protein yield were obtained, and the interaction between the five factors was investigated in this study. In the finding for the lipase concentration, $(NH_4)_2SO_4$ and K_2HPO_4 were significant contributing factors, while only K_2HPO_4 was found to contribute to the total protein. However, the Pareto plot displayed that the K_2HPO_4 and $(NH_4)_2SO_4$ have different influences on lipase production. Meanwhile, K_2HPO_4 showed a positive impact on both lipase and total protein concentration. In the study, the maximum lipase concentration was 5.56 mg/ml with lipase activity 7.58 x 10^{-2} µmol/min/ml and maximum specific lipase activity was 3.01 x 10^{-2} µmol/min/mg. While for the total protein, the highest concentration was 1.28 mg/ml.

Keywords: fractional factorial design**,** recombinant *E. coli*, recombinant lipase, two-level fractional factorial design

Article History: - Received: 20 November 2020; Accepted: 20 August 2021; Published: 31 October 2021

Introduction

Lipases or its systematic name, triacylglycerol acylhydrolase (EC 3.1.1.3), fit in hydrolases class and water-soluble enzymes that catalyse the hydrolysis of insoluble triacylglycerol to produce free fatty acids, diacylglycerol, monoacylglycerol and glycerol (Geoffry and Achur, 2018). Hence, this enzyme is one of the essential classes of industrial enzymes in basic and applied research (Kaur, 2017). Lipases are usually obtained from various species of plants, animals, bacteria, fungi, and yeast. However, lipases produced by bacteria are the most beneficial and commercially more critical compared to the ones made by plants and animals due to their stability and the ease of their cultivation and optimisation to gain higher yield (Bharathi et al., 2019). Lipases are used in various areas, mainly as a biological catalyst to produce products such as food ingredients and their application in making fine chemicals. Lipases are also used in detergents, foods, oils, and fats (including biodiesel), leather, pulp and paper, textiles, and animal feed. A large amount of lipase usage but lack of its mass production in the food processing industry has also contributed to this research's idea (Ray, 2012). Besides, the study had changed the biochemistry area during the periods where the purification and extraction of naturally produced protein

from animal and plant tissues are almost limited (Rosano and Ceccarelli, 2014). Researchers back then and now are still studying the most effective and efficient ways to get the purified recombinant protein.

Furthermore, according to the Markets (2021) report, the market for industrial enzymes is expected to increase at a compound annual growth rate (CAGR) of 6.4% over the estimated 2021 to 2028. Microbial lipases had the largest share of 61.64% in 2020, according to the report. The availability of genetic therapy, such as recombinant lipases, is one of the causes driving up microbial lipase synthesis. Due to rising demand for dairy products and the region's burgeoning pharmaceutical industry, Asia Pacific will account for the most significant part of worldwide region markets in 2020, with 33.8% (Markets, 2021). In conclusion, regardless of the natural or recombinant form, the demand for lipase is massive.

Due to the recombinant lipase markets, the use of a recombinant host is equally important. There are so many advantages of *E. coli* that make it the preferable host cell in recombinant proteins. The main advantage of *E. coli* is that it is a fast-growing organism when treated with suitable and optimal environmental conditions (Ahmad et al., 2018). Besides, low-cost complex media that can be made from available components can be used for their growth. Moreover, high cell density can be easily achieved where *E. coli* is boosted even while producing the recombinant protein (Rosano and Ceccarelli, 2014). The manufacturing of recombinant protein using *E. coli* has been broadly studied since the modern biotechnology era in the mid and late 70s (Makino et al., 2011).

Besides the recombinant host, the cultivation medium is also crucial to support the production of recombinant lipase. Thus, in achieving screening and optimisation of medium composition and conditions of a fermentation process, many approaches of experiment design can be conducted. In early research, most researchers have used one-factor-at-a-time (OFAT). Despite the success of using OFAT, Hu et al., (2015) in certain aspects, it still suffers from inconclusive interaction between the investigated factors and needed support from the statistical design such as response surface design (RSM). Thus, statistical design in many microbial productions is highly recommended (Nor et al., 2017). Under RSM, experimental designs such as central composite design (CCD) and factorial design (FD) can combine factors concurrently. In statistical design, more than three factors can be investigated simultaneously, and the contributing factors can be analysed according to the hierarchy of the factors to explain the response. Furthermore, the experimental and anticipated results may be compared using statistical design, and the most relevant factors between components can be investigated (Nor et al., 2017).

Besides being convenient in the analysis and number of experiments due to constraints of time and resources, the statistical design can use a portion of the full FD, thus converting the full FD to fractional FD (FFD). A FFD is an orthogonal array design that allows experimenters to explore primary and desired interaction effects in a limited number of trials or experimental runs. These fractional factorial designs are the most popular and widely used types in many research and industry (Antony, 2014). Additionally, statistical design can save the time and cost of the experimental works when many independent variables are assessed (Akbari et al., 2010; Watkins and Newbold, 2020). The FFD design is based on the level, factor, and fraction index (Antony, 2014). The common FFD is two levels of parameters for each factor (k) and the fraction index (p); thus, the general notation for two-level FFD is 2^{k-p}.

In this study, five factors of medium components $(NH_4)_2SO_4$, glycerol, yeast extract, KH_2PO_4 and K_2HPO_4 with two levels of parameters (low and high range) were investigated, giving full FD of 32 experiments (2^5). Due to time constraints, a fractional design is chosen. Thus, the number of FFD is reduced to half of the total experiment, which is 16 (2^{5-1}). The importance of this design is to screen which factors significantly influence the production of lipase from the recombinant *E. coli*. Design Expert Software generated the FFD by exploiting the medium composition and parameter ranges. The research is inspired by similar work in the area. Based on a prior study conducted by Akbari et al., (2010), it was determined that this approach was practical, as just a portion number of tests were utilised to screen various critical aspects that could influence the output. In particular, the current study is interested in determining the effect of the medium composition of r*E. coli* producing recombinant lipase and total protein using the two-level FFD.

Methods

Microorganism, Growth Medium and Conditions

The r*E. coli* containing lipase gene from *Staphylococcus hyicus* was obtained from Universiti Putra Malaysia. The r*E. coli* inoculum was prepared using Luria Bertani, 20 g/L in 250 mL Erlenmeyer flasks with a working volume of 50 mL. The medium was autoclaved at 121 °C for 15 min. Sterilised filtered ampicillin with a final concentration of 50 µg/mL was added to the inoculum at 0 h. The inoculum was agitated at 10 rpm, 30 °C for 12 h (Omar et al., 2017).

Two-level Fractional Factorial Design

Design Expert Software version 7.1.6 (Minnesota, USA) was used to design the experiment and analyse this study's output. Based on the basal fermentation medium, five medium components are included in the experimental design, whereas the remaining component was a constant. The respective components were coded as A - $(NH_4)_2SO_4$), B - glycerol, C - yeast extract, D - KH_2PO_4 and E - K_2HPO_4 with low and high ranges and were tabulated in Table 1.

Table 1. The coded factors and their respective low and high ranges.

Factors	Unit	Low level	High level
A-$(NH_4)_2SO_4$	g/L	7.0	14.0
B-Glycerol	g/L	17.5	35.0
C-Yeast extract	g/L	10.0	20.0
D-KH_2PO_4	g/L	1.0	2.0
E-K_2HPO_4	g/L	8.25	16.5

A total of 16 run numbers, equal to the number of submerged batch fermentation in shake flasks, as shown in Table 2, was conducted in this study. The medium then was adjusted to pH 7 before autoclaving. The glycerol was prepared separately and added aseptically into the culture medium before inoculation. To maintain the plasmid that carries the recombinant genes, sterilised filtered ampicillin was added until the final 50 µg/mL (Omar et al., 2017). After that, 10 % (v/v) of inoculum with OD_{600nm} ~ 0.7 was added, and the culture was shaken at 100 rpm, 30 °C. The culture was incubated for 3 h and later induced with a final concentration of 1 mM IPTG. The induced cultured was continued to set for another 6 h. After a total of 9 h of incubation, the culture was transferred into sterilised 50 mL centrifuge tubes and kept at 4 °C for 24 hours before further analysis can be done.

Table 2. Experimental design with the coded factors.

Run	Factor A: $(NH_4)_2SO_4$ (g/50mL)	Factor B: Glycerol (g/50mL)	Factor C: Yeast extract (g/50mL)	Factor D: KH_2PO_4 (g/50mL)	Factor E: K_2HPO_4 (g/50mL)
1	0.7	0.875	1	0.05	0.825
2	0.7	0.875	0.5	0.1	0.825
3	0.35	1.75	1	0.1	0.4125
4	0.35	0.875	0.5	0.1	0.4125
5	0.35	1.75	0.5	0.05	0.4125
6	0.35	1.75	0.5	0.1	0.825
7	0.7	1.75	0.5	0.05	0.825
8	0.7	0.875	0.5	0.05	0.4125
9	0.7	0.875	1	0.1	0.4125
10	0.35	0.875	1	0.1	0.825
11	0.7	1.75	1	0.1	0.825
12	0.7	1.75	1	0.05	0.4125
13	0.35	0.875	1	0.05	0.4125
14	0.35	1.75	1	0.05	0.825
15	0.7	1.75	0.5	0.1	0.4125
16	0.35	0.875	0.5	0.05	0.825

Lipase assay

Before the lipase assay, the cells were lysed using glass-bead and lysis buffer to harvest the intracellular protein. Lipase concentration was measured using spectrophotometry with the addition of substrate, p-nitrophenyl palmitate (pNPP, Sigma-Aldrich). As for the reaction mixture, 850 µL of solution A containing 50 mM Tris-HCl buffer pH 8, 0.1 % (w/v) gum Arabic and 0.2 % (w/v) sodium deoxycholate and 50 µL of solution B containing 10 mM pNPP in isopropanol were carefully prepared. Before the assay, 100 µL of the sample was obtained and was incubated for 1 h. The sample then was mixed with a reaction mixture and incubated at 35 °C for 6 min in a water bath. Next, 500 µL of 3 M of HCl was added into the mix and centrifuged at 10 000 rpm for 1 min. The 500 µL of clear supernatant was transferred to a cuvette, and next, 1 mL of 2 M of NaOH was added. The mixture was read at 410 nm by spectrophotometer against a free enzyme mixture as control or blank (Omar et al., 2017). A standard curve of pure lipase from *Rhizopus oryzae* (Sigma Aldrich) was prepared using a similar assay with different concentrations from 1 to 10 mg/ml. The amount of 4-nitrophenol produced by the reaction was compared between the standard using the extinction coefficient for the samples.

Total protein assay

The total protein was determined by the Lowry method. The assay was performed using a Protein Assay Lowry kit (Nacalai Tesque, Inc.). The sample solution was measured at absorbance 750 nm using a spectrophotometer. The Bovine Serum Albumin (BSA) was used at different concentrations from 0.2 to 1 mg/ml as the standard solution to compare the sample's concentration. The standard equation showed, one absorbance was equal to 0.98 mg/ml of BSA.

Result and Discussion

The factors used in this study were selected based on the previous research by Gu et al., (2016) and Islam et al., (2007). Previous studies had shown the effectiveness of each factor used to produce and determine lipase and total protein from r*E. coli* cells. For the statistical analysis, ANOVA has been performed to determine the significant difference of *p*-value at <0.05. The major responses were lipase concentration and total protein was analysed, respectively. Besides ANOVA, perturbation graphs and Pareto plots were included to display the positive and negative effects of the factors on the responses. Pareto analysis is one of the approaches determining the most influential factors and any combination of factors based on the Bonferroni limit and t-value limit. Factors that reached the Bonferroni limit are the statistically significant factors. In contrast, the factors that reached the t-value limit are the factors that are likely to be statistically significant (Kikwai et al., 2012).

Effect of medium compositions on lipase concentration

E. coli is not a lipase producer. Thus, the lipase reaction to the substrate was assumed solely from the recombinant protein induced by the IPTG. Even though the incubation was considerably short in this study compared to 12 h by Batumalaie et al., (2017) and 40 h by Zhang et al., (2018), the induced r*E. coli* was able to produce lipase fast, similarly to Kaur (2017), which was 3 h. The highest lipase concentration obtained was 5.56 mg/ml from run number 6, shown in Table 2. Extension to the amount of the lipase, the maximum lipase activity was 7.58×10^{-2} µmol/min/ml, and maximum specific lipase activity was 3.01×10^{-2} µmol/min/mg.

From Table 3, the ANOVA shows that the model of lipase concentration was found to be significant. The two main factors were also significant, $(NH_4)_2SO_4$, *p*-value = 0.0005 and K_2HPO_4, p-value = <0.0001; whereas all combinations of two-factor interactions, were significant. This result shows that the proposed experimental design is at least as effective as the more common methods and able to recognise the significant interactions between the factors (Nor et al., 2017). The $(NH_4)_2SO_4$ was surprisingly contributing more impact to the lipase production rather than yeast extract, and this finding contrasts with the data reported by Chai and Adnan (2018) and Jørgensen (2009). However, $(NH_4)_2SO_4$ is recommended substitution for nitrogen supply in the fermentation as yeast extract is more expensive and harder to accommodate the specific need for the *E. coli* maintenance and growth (Ginésy et al., 2017). The addition of yeast extract as the nitrogen source in the experimental medium makes it rich in nutrients and may promote the bacteria's growth (Gu et al., 2016; Shafqat et al., 2015).

Table 3. ANOVA for lipase concentration.

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	50.43	11	4.58	82.12	0.0003
A	5.83	1	5.83	104.47	0.0005
B	0.034	1	0.034	0.62	0.4761
C	0.022	1	0.022	0.40	0.5632
D	0.037	1	0.037	0.67	0.4600
E	29.02	1	29.02	519.78	< 0.0001
AC	2.16	1	2.16	38.77	0.0034
AE	0.95	1	0.95	17.10	0.0144
BC	1.85	1	1.85	33.16	0.0045
BD	2.51	1	2.51	44.96	0.0026
BE	0.46	1	0.46	8.21	0.0456
DE	7.55	1	7.55	135.16	0.0003
Residual	0.22	4	0.056		
Corrected Total	50.66	15			
$\mathbf{R^2}$					0.9956
Adjusted $\mathbf{R^2}$					0.9835
Adequate Precision					29.005
Coefficient of Variance %					5.62

A-$(NH_4)_2SO_4$, B-Glycerol, C-Yeast extract, D-KH_2PO_4, E-K_2HPO_4, the significant value of $p < 0.05$

The regression coefficient was used to assess the goodness of fit (R^2). In this case, the regression coefficient ($R^2 = 0.9956$) revealed that the proposed regression model was not explained at just 0.44 % of the overall variations. Furthermore, the difference between R^2 and adjusted R^2 was less than 0.2, indicating that the model is reliable for interpolation. Additionally, the low value of the coefficient of variation (CV=5.62%) demonstrated a high degree of reliability and precision of the conducted experiments (Moorthy et al., 2015). The signal-to-noise ratio is measured by adequate precision, and a ratio larger than 4 is desirable for navigating design space. In this situation, the appropriate precision was determined to be 29.01, indicating the best fitness of the created model. Adequacy necessitates knowledge on the residuals' lack of fit. Model diagnostic charts, such as the predicted value against the experimental value graph, aided in displaying the link between the experimental and predicted values and evaluating model sufficiency. Ensuring that the fitted linear model gives a broad approximation of the actual values while ignoring small, deceptive screening and optimisation impacts is necessary. The data points in the graph (Figure 1A) drawn between the predicted and actual values were found to be adjacently spaced, indicating the minimal variation and efficacy concordance between the expected and actual values. An internally studentised residual versus experimental runs plot was created to check that the generated model performed satisfactorily. In the residual versus run plot (Figure 1B), a random trend was detected. All data points fell within the range of control limits, suggesting that the experiments were carried out randomly, reducing the possibility of mistakes assuring adequate fit (Teja and Damodharan, 2018).

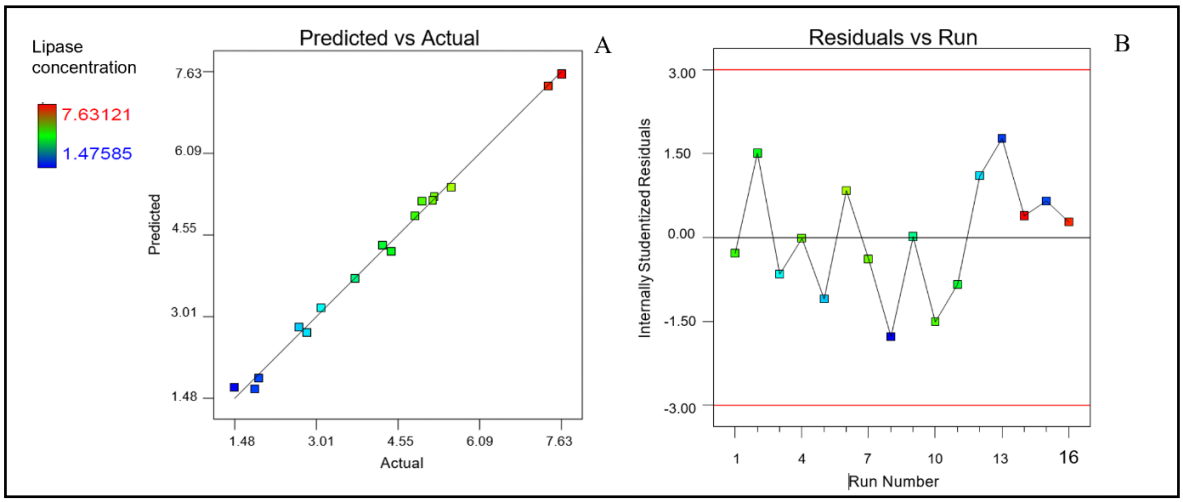

Figure 1. The diagnostics plots of lipase concentration. A. Predicted versus actual data, B. Internally studentised residuals versus the run number.

The interactions between factors of medium compositions were observed in Figure 2. Based on Figure 2A, the perpendicular lines clearly show the interaction of low level and high levels of the yeast extract and glycerol. The same pattern in Figure 2B and 2C show the strong interaction between yeast extract and $(NH_4)_2SO_4$ and KH_2PO_4, and glycerol, respectively. Therefore, the three interactions between the two-factor are contributing to the yield of lipase concentration. The remaining two-factor, however, display weak interaction and does not display differences at any level. Hence, it suggests a lack of interactions between the factors (Bolton and Bon, 2004).

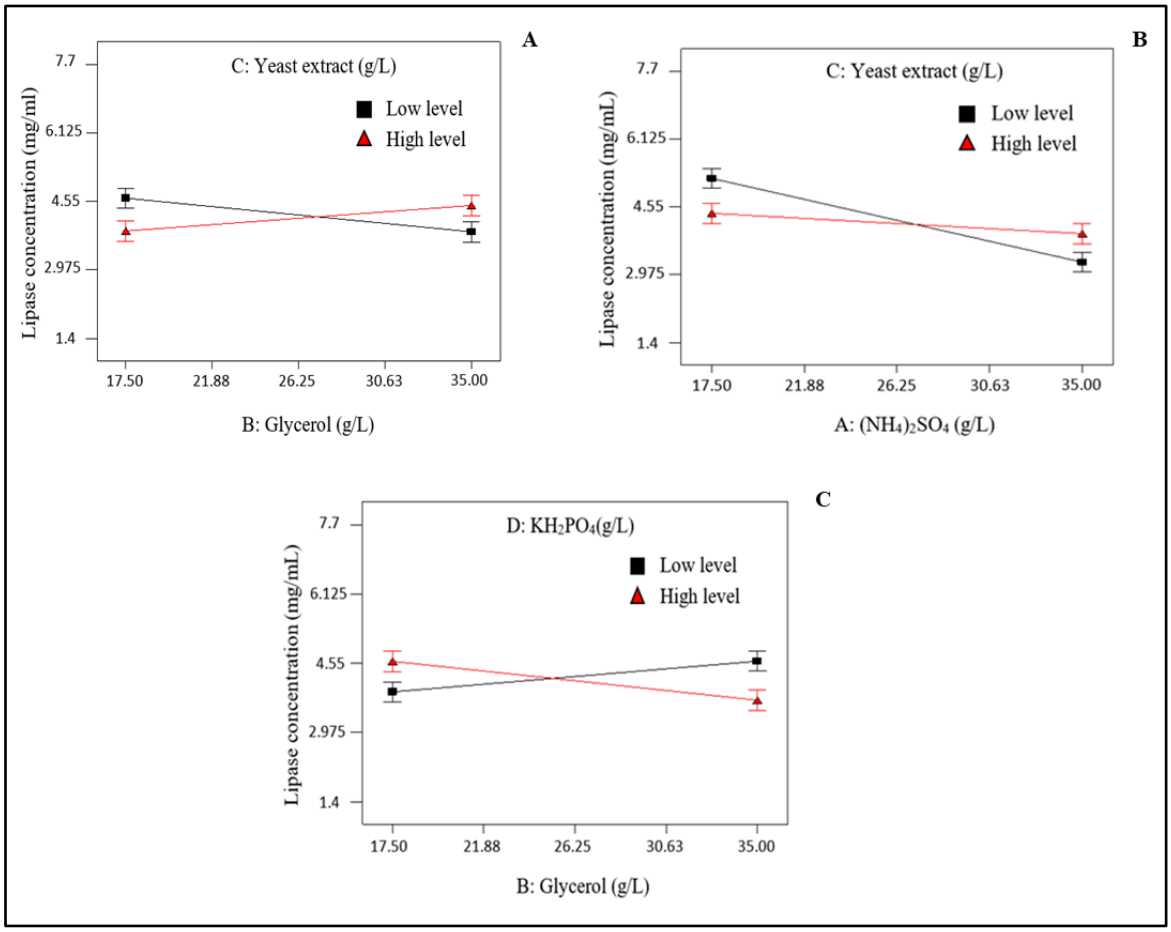

Figure 2. Interaction plots of the selected two-factor interactions. A. Glycerol and yeast extract; B. $(NH_4)_2SO_4$ and yeast extract; C. Glycerol and KH_2PO_4.

The Pareto plots were used to illustrate the importance of the factors and the response coefficient significance study on the yield of lipase and total protein in Figures 3 and 5, respectively. Referring to the Pareto charts, blue bars represent a negative effect. In contrast, orange bars represent a positive effect, while the main factors or two-factors interactions were displayed on the top of the bar. The positive impact of the factors suggests that the presence of the factors is most likely to influence the response. The negative effects of the factors mean the presence of the factors is less likely to affect the production.

For lipase yield, the Pareto chart (Figure 3) shows K_2HPO_4 contributed the most statistical significance, followed by the combination of KH_2PO_4 and K_2HPO_4, a single factor of $(NH_4)_2SO_4$, a combination of glycerol and KH_2PO_4, the combination of $(NH_4)_2SO_4$ and yeast extract. The mixture of glycerol and yeast extract also greatly affected the statistical significance of the lipase concentration. Here, the influence of significant factor in ANOVA Table 3 explained that $(NH_4)_2SO_4$ has shown a negative effect on the lipase concentration. Still, in combination with yeast extract, it was a positive effect. Thus, indicating that having a rich medium such as yeast extract supported any lacking nutrient of a defined component such as $(NH_4)_2SO_4$ in the lipase production.

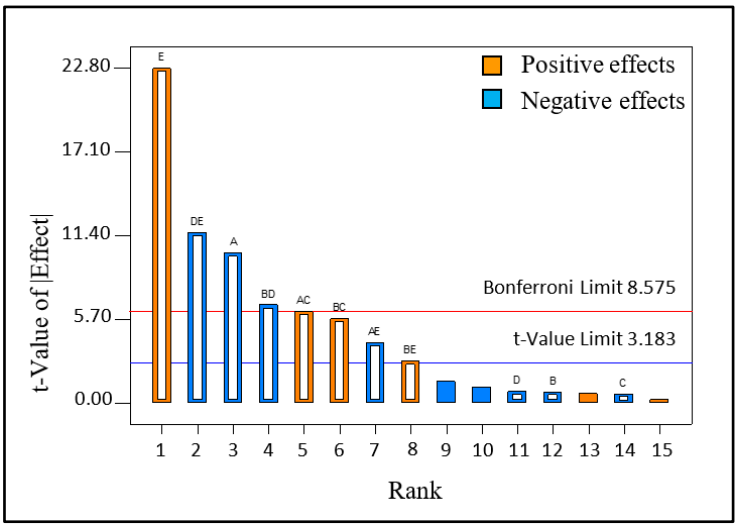

Figure 3. Response coefficient significance study on lipase concentration.

Effect of medium compositions on total protein concentration

The study on the total protein of the cells was conducted through the total protein assay procedure. The maximum protein content was 1.28 mg/ml from run number 7 by Table 2. The ANOVA in Table 4 indicates that the model for total protein and main factor, K_2HPO_4, is significant at *p*-value = 0.0275 and *p*-value = 0.0024, respectively. Likewise, all terms except the combination of $(NH_4)_2SO_4$ and yeast extract; *p*-value = 0.0327 does not show any contribution in the two-factor interactions. The evaluation suggests that the obtained results are accurate to the expectations of this study.

For total protein, the regression coefficient (R^2 = 0.9817) revealed that the proposed regression model was not explained at 1.83% of the overall variations, a bit higher than the lipase concentration model in Table 3. Furthermore, the difference between R^2 and adjusted R^2 for the total protein concentration was less than 0.2, indicating that the model is acceptable for interpolation. However, the low value of the coefficient of variation (CV = 11.92%) demonstrated a high degree of reliability and precision of the conducted experiments. The signal-to-noise ratio is measured by adequate precision, and a ratio larger than 4 is desirable for navigating design space. In this study, the adequate precision was determined to be 11.527, indicating an adequate signal to navigate the design space. Similar to lipase concentration, model diagnostic charts were provided in evaluating model sufficiency for total protein. The data points in the graph (Figure 4A) drawn between the predicted and actual values were found to be scattered along the line, indicating there is variation between the expected and actual values. Nevertheless, a random trend was detected in the residual versus run plot (Figure 4B). All data points fell within the

range of control limits, suggesting that the experiments were carried out in a random trend, thus assuring a good fit of the model.

Table 4. ANOVA for the total protein.

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	1.22	12	0.10	13.40	0.0275
A	0.018	1	0.018	2.34	0.2237
B	0.040	1	0.040	5.21	0.1067
C	0.067	1	0.067	8.80	0.0592
D	0.019	1	0.019	2.49	0.2127
E	0.70	1	0.70	92.78	0.0024
AB	0.018	1	0.018	2.35	0.2226
AC	0.11	1	0.11	14.20	0.0327
BC	0.058	1	0.058	7.60	0.0703
BE	0.061	1	0.061	8.10	0.0653
CD	0.049	1	0.049	6.47	0.0844
CE	0.054	1	0.054	7.12	0.0758
DE	0.025	1	0.025	3.34	0.1652
Residual	0.023	3	7.591E-003		
Corrected Total	1.24	15			
$\mathbf{R^2}$					0.9817
Adjusted $\mathbf{R^2}$					0.9084
Adequate Precision					11.527
Coefficient of Variance %					11.92

A-$(NH_4)_2SO_4$, B-Glycerol, C-Yeast extract, D-KH_2PO_4, E-K_2HPO_4, the significant value of $p < 0.05$

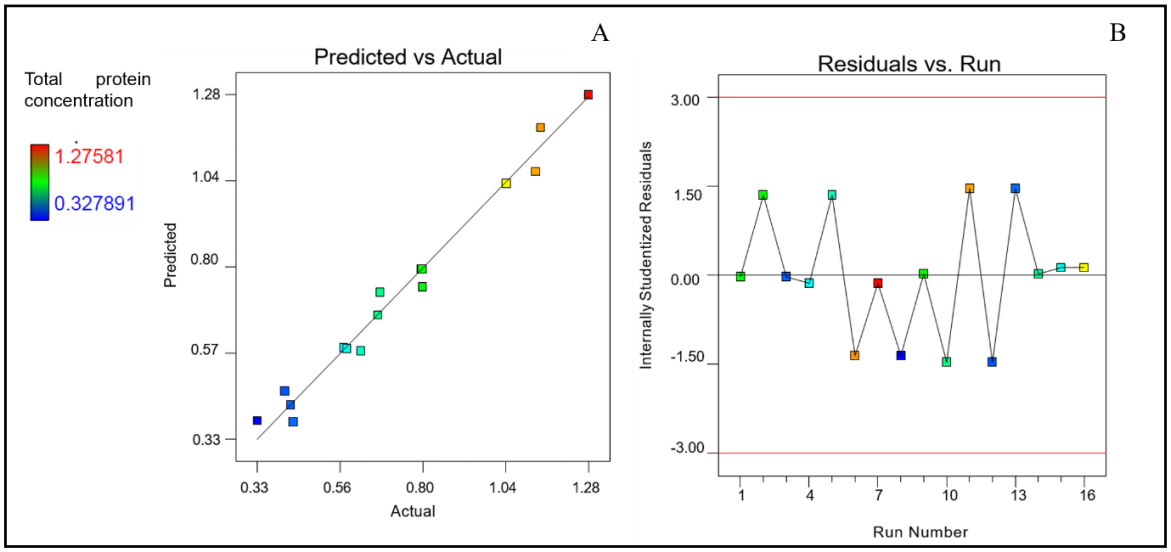

Figure 4. The diagnostics plots of total protein concentration. A. Predicted versus actual data, B. Internally studentised residuals versus the run number.

Figure 5 displays the interactions by observing the perpendicular lines of low and high levels between yeast extract and $(NH_4)_2SO_4$. It was indicated that the interactions for these factors have weak significance to the yield of total protein. The remaining interactions do not show any connections at any level. In other words, there is a lack of interaction between the medium components for maximising the total protein.

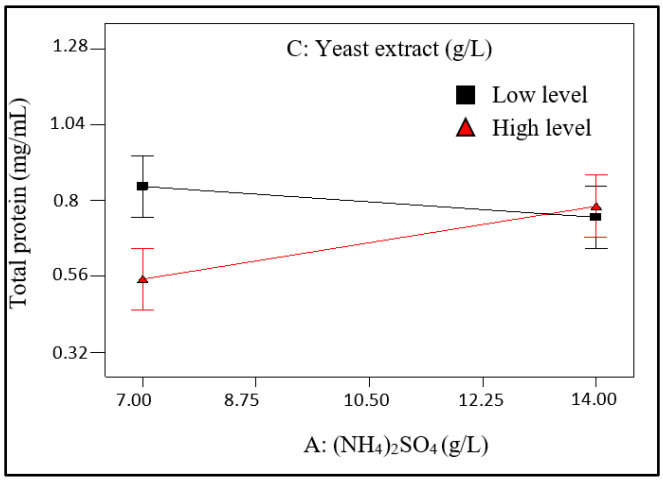

Figure 5. Interaction between $(NH_4)_2SO_4$ and yeast extract.

In agreement with lipase concentration, K_2HPO_4 also positively contributed to the total protein in Pareto analysis (Figure 6). This finding is interesting, and it could be hypothesised that the production of lipase and total protein in r*E. coli* is more suitable to use K_2HPO_4 rather than KH_2PO_4 for this study. Both medium components are the provider for the phosphorus component in the ATP for microbial growth. This finding agrees with an investigation on the medium composition to other microbial such as *Streptomyces venezuelae* by Yi et al., (2015). However, a contrast reported by Sugumaran et al., (2013) showed that *Aureobasidium pullulans* preferred KH_2PO_4 instead of K_2HPO_4. The significant amount used for both factors provide some buffering ability to the medium and may limit the reduction in pH throughout the fermentation process (Islam et al., 2007). Thus, it is difficult to draw any conclusions based on the results shown here.

The combination of $(NH_4)_2SO_4$ with yeast extract had fairly passed the t-value limit to be statistically significant. Furthermore, other factors in the total protein procedure did not reach both the Bonferroni limit and the t-value limit, thus not influencing the total protein production by the r*E.coli.*

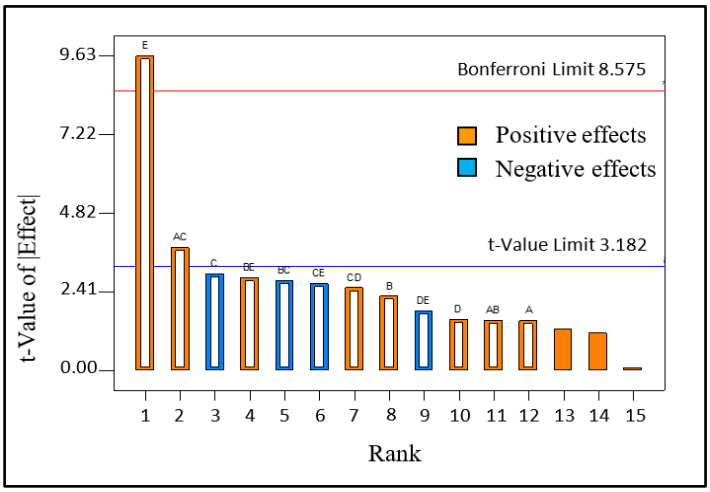

Figure 6. Response coefficient significance study on total protein.

Limitation of the study

The addition of IPTG in the experiment plays a vital role in ensuring recombinant protein production while slowing down bacterial growth. The main limitation, however, is the lack of time to conduct the experiments. The incubation time after the induction of IPTG also affected the amount of lipase produced by the r*E. coli.* In this study, only six hours of incubation were managed for intracellular lipase production in the r*E. coli.* However, the lipase assay is sensitive and adequate to prove that the lipase was from the recombinant r*E. coli*.

Conclusion

The statistical design can explain the interaction of two-factor interactions using ANOVA analysis, perturbation plot, Pareto chart, and diagnostics plots to justify the chosen factors. The current study used an experimental design approach, 2^{5-1} fractional factorial design, to screen the most influential medium compositions by a recombinant *E. coli* to produce lipase and total protein in batch fermentation. The results show that the experimental design able to identify that K_2HPO_4 had contributed to provide the most positive effect in both responses and reached above the Bonferroni limit. For lipase production, $(NH_4)_2SO_4$ was found significant and negative influence as a sole factor. However, combination with yeast extract and $(NH_4)_2SO_4$ was likely to be a positive influence. In future experimental designs for screening or optimisation, factors such as glycerol and K_2HPO_4 can be kept constant, and other factors such as incubation time, agitation rate, and pH can be considered.

Acknowledgement

The authors are thankful to Associate Professor Dr Adam Thean Chor from Universiti Putra Malaysia for the recombinant lipase culture. This research uses minimum resources from the previous research fund under the Ministry of Higher Education (MOHE) coded as FRGS/1/2016/TK02/UMP/02/15.FRGS/1/2016/TK (RDU160122). The authors would like to thank the assistant science officer of the biology laboratory at the Universiti Teknologi MARA Cawangan Pahang, Kampus Jengka, for the help in some of the laboratory activities.

References

Ahmad, I., Nawaz, N., Darwesh, N. M., Ur Rahman, S., Mustafa, M. Z., Khan, S. B., & Patching, S. G. (2018). Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*. *Protein Expr Purif, 144*, 12-18. doi:10.1016/j.pep.2017.11.005

Akbari, N., Khajeh, K., Ghaemi, N., & Salemi, Z. (2010). Efficient refolding of recombinant lipase from *Escherichia coli* inclusion bodies by response surface methodology. *Protein Expr Purif, 70*(2), 254-259. doi:10.1016/j.pep.2009.10.009

Antony, J. (2014). *Design of Experiments for Engineers and Scientists* (Second Edition ed.). USA: Elsevier Ltd. .

Batumalaie, K., Khalili, E., Mahat, N. A., Huyop, F. Z., & Wahab, R. A. (2017). A statistical approach for optimizing the protocol for overexpressing lipase KV1 in *Escherichia coli*: purification and characterization. *Biotechnology & Biotechnological Equipment, 32*(1), 69-87. doi:10.1080/13102818.2017.1407670

Bharathi, D., Rajalakshmi, G., & Komathi, S. (2019). Optimization and production of lipase enzyme from bacterial strains isolated from petrol spilled soil. *Journal of King Saud University - Science, 31*(4), 898-901. doi:10.1016/j.jksus.2017.12.018

Bolton, S., & Bon, C. (2004). *Pharmaceutical Statistics : Practical and Clinical Applications* (Vol. 135). New York, USA: Marcel Dekker, Inc.

Chai, J. M., & Adnan, A. (2018). Effect of different nitrogen source combinations on microbial cellulose production by *Pseudomonas aeruginosa* in batch fermentation. *IOP Conf. Series: Materials Science and Engineering, 440*(012044), 1-9. doi:doi:10.1088/1757-899X/440/1/012044

Geoffry, K., & Achur, R. N. (2018). Screening and production of lipase from fungal organisms. *Biocatalysis and Agricultural Biotechnology, 14*, 241-253. doi:10.1016/j.bcab.2018.03.009

Ginésy, M., Rusanova-Naydenova, D., & Rova, U. (2017). Tuning of the carbon-to-nitrogen ratio for the production of l-arginine by *Escherichia coli*. *Fermentation, 3*(4). doi:10.3390/fermentation3040060

Gu, Y.-Q., Zhang, X.-H., Li, B.-B., & Chen, W.-Q. (2016). *Screening and molecular identification of a bacterial strain with higher lipase productivity.* Paper presented at the 2016 Joint International Conference on Social Science and Environmental Science (SSES 2016) and International Conference on Food Science and Engineering (ICFSE 2016).

Hu, Y., Qin, H., Zhan, Z., Dun, Y., Zhou, Y., Peng, N., Ling, H., Liang, Y., & Zhao, S. (2015). Optimization of

Saccharomyces boulardii production in solid-state fermentation with response surface methodology. *Biotechnology & Biotechnological Equipment, 30*(1), 173-179. doi:10.1080/13102818.2015.1086689

Islam, R. S., Tisi, D., Levy, M. S., & Lye, G. J. (2007). Framework for the rapid optimization of soluble protein expression in *Escherichia coli* combining microscale experiments and statistical experimental design. *Biotechnol. Prog., 23*, 785−793.

Jørgensen, H. (2009). Effect of nutrients on fermentation of pretreated wheat straw at very high dry matter content by *Saccharomyces cerevisiae*. *Applied Biochemistry and Biotechnology, 153*, 44–57.

Kaur, J. (2017). Studies on recombinant lipase production by *E. coli*: Effect of media and bacterial expression system optimization. *International Journal of Molecular Biology, 2*(1). doi:10.15406/ijmboa.2017.02.00008

Kikwai, L., Tran, D., Hauck, W. W., Shah, V. P., & Stippler, E. S. (2012). Effect of various operational parameters on drug release from a 1% hydrocortisone semisolid dosage form using the vertical diffusion cell apparatus. *Dissolution Technologies, 19*(3), 6-13. doi:10.14227/dt190312p6

Makino, T., Skretas, G., & Georgiou, G. (2011). Strain engineering for improved expression of recombinant proteins in bacteria. *Microbial Cell Factories, 10*(32), 1-10.

Markets, F. (2021). Lipase market by source, application, region, global industry analysis, market size, share, growth, trends, and forecast 2021 to 2028 (Market Report). (419242). Retrieved July 2021, from Fior Markets https://www.fiormarkets.com/report/lipase-market-by-source-animal-lipases-microbial-lipases-419242.html

Moorthy, I. G., Maran, J. P., Surya, S. M., Naganyashree, S., & Shivamathi, C. S. (2015). Response surface optimization of ultrasound assisted extraction of pectin from pomegranate peel. *Int J Biol Macromol, 72*, 1323-1328. doi:10.1016/j.ijbiomac.2014.10.037

Nor, N. M., Mohamed, M. S., Loh, T. C., Foo, H. L., Rahim, R. A., Tan, J. S., & Mohamad, R. (2017). Comparative analyses on medium optimization using one-factor-at-a-time, response surface methodology, and artificial neural network for lysine–methionine biosynthesis by *Pediococcus pentosaceus* RF-1. *Biotechnology & Biotechnological Equipment, 31*(5), 935-947. doi:10.1080/13102818.2017.1335177

Omar, W. S. A. W., Sulaiman, A. Z., Ajit, A., Chisti, Y., & Chor, A. L. T. (2017). *Effect of flow rate, duty cycle, amplitude, and treatment time of ultrasonic regimens towards Escherichia coli harbouring lipase.* Paper presented at the 29th Symposium of Malaysian Chemical Engineers (SOMChE) 2016 Miri, Sarawak, Malaysia.

Ray, A. (2012). Application of lipase in industry. *Asian J. Pharm. Tech., 2*(2), 33-37.

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol, 5*, 172. doi:10.3389/fmicb.2014.00172

Shafqat, I., Shahzad, S., & Yasmin, A. (2015). Microbial potential of lipase production from different industrial effluents. *Journal of Applied Biological Sciences, 9*(1), 15-19.

Sugumaran, K. R., Sindhu, R. V., Sukanya, S., Aiswarya, N., & Ponnusami, V. (2013). Statistical studies on high molecular weight pullulan production in solid state fermentation using jack fruit seed. *Carbohydr Polym, 98*(1), 854-860. doi:10.1016/j.carbpol.2013.06.071

Teja, S. P. S., & Damodharan, N. (2018). 2^3 full factorial model for particle size optimization of methotrexate loaded chitosan nanocarriers: a design of experiments (DoE) approach. *Biomed Res Int, 2018*, 7834159. doi:10.1155/2018/7834159

Watkins, E. R., & Newbold, A. (2020). Factorial designs help to understand how psychological therapy works. *Front Psychiatry, 11*, 429. doi:10.3389/fpsyt.2020.00429

Yi, J. S., Kim, M. S., Kim, S. J., & Kim, B. G. (2015). Effects of sucrose, phosphate, and calcium carbonate on the production of pikromycin from *Streptomyces venezuelae*. *J Microbiol Biotechnol, 25*(4), 496-502.

doi:10.4014/jmb.1409.09009

Zhang, W., Lu, J., Zhang, S., Liu, L., Pang, X., & Lv, J. (2018). Development an effective system to expression recombinant protein in *E. coli* via comparison and optimization of signal peptides: Expression of *Pseudomonas fluorescens* BJ-10 thermostable lipase as case study. *Microb Cell Fact, 17*(1), 50. doi:10.1186/s12934-018-0894-y